

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2001, 18:06:14 ; Search time 2667.65 Seconds  
(without alignments)  
4145.761 Million cell updates/sec

Title: US-09-784-340-3\_COPY\_197\_911  
Perfect score: 715  
Sequence: 1 atgaggtctgcaagtcagc.....gtttatagtaagcattag 715

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues  
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_p11: \*  
13: gb\_p12: \*  
14: gb\_p13: \*  
15: gb\_p14: \*  
16: em\_ba1: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_rod: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_or: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_p1: \*  
48: em\_ro: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_v1: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_v11: \*  
59: gb\_v12: \*  
60: gb\_htg1: \*  
61: gb\_htg2: \*  
62: gb\_htg3: \*  
63: gb\_htg4: \*  
64: gb\_htg5: \*  
65: gb\_htg6: \*  
66: gb\_htg7: \*  
67: gb\_htg8: \*  
68: gb\_htg9: \*  
69: gb\_htg10: \*  
70: gb\_htg11: \*  
71: gb\_htg12: \*  
72: gb\_htg13: \*  
73: gb\_htg14: \*  
74: gb\_htg15: \*  
75: gb\_htg16: \*  
76: gb\_htg17: \*  
77: gb\_htg18: \*  
78: gb\_htg19: \*  
79: gb\_htg20: \*  
80: gb\_htg21: \*  
81: gb\_htg22: \*  
82: gb\_htg23: \*  
83: gb\_htg24: \*  
84: gb\_htg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_ro1: \*  
95: gb\_ro2: \*  
96: gb\_in4: \*  
97: gb\_pr10: \*  
98: em\_ba3: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	703.8	98.4	212904	66	AC021146	AC021146 Homo sapi
2	539.8	75.5	2786	89	AK025587	AK025587 Homo sapi
3	407.8	57.0	139015	69	AC025488	AC025488 Homo sapi
4	407.8	57.0	169246	62	AC011254	AC011254 Homo sapi
C 5	406	56.8	138232	63	AC012582	AC012582 Homo sapi
C 6	391.6	54.8	226077	63	AC013296	AC013296 Homo sapi
7	348.8	48.8	2547	94	AF175221	AF175221 Cavia por
8	190.2	26.6	2090	97	HS008854	HS008854 Human UDP g

9	190.2	26.6	181117	87	AC019173	AC019173 Homo sapi
10	188.6	26.4	2114	89	AF180322	AF180322 Homo sapi
11	188.4	26.3	2091	88	AF081793	AF081793 Homo sapi
12	188.4	26.3	2108	89	AF112112	AF112112 Macaca fa
13	186.8	26.1	2079	93	HSASJ162	HSASJ162 Homo sapi
14	186.8	26.1	2093	97	HSUDGT	HSUDGT Human mRNA
15	186.8	26.1	2097	88	AF064200	AF064200 Homo sapi
16	186.8	26.1	27310	89	AF135416	AF135416 Homo sapi
17	186.2	25.6	1650	9	AX040087	AX040087 Sequence
18	183.2	25.6	1671	7	RABUG2B	L01082 Oryctolagus
19	182.2	25.5	2123	97	HSU06641	U06641 Human UDP g
20	179	25.0	2107	97	HSU59209	U59209 Homo sapien
21	175.8	24.6	2075	88	AF072223	AF072223 Macaca fa
22	172.4	24.1	1855	97	H0MUDPGTA	J05428 Human 3/4-c
23	171.4	24.0	2634	95	RATUDPGT	J02589 Rat UDP glu
24	171.4	24.0	11728	95	RATUDPGV	M7439 Rattus ratt
25	169.2	23.7	1662	89	AF294902	AF294902 Macaca mu
26	169.2	23.7	1722	88	AF016492	AF016492 Homo sapi
27	166.6	23.3	1832	7	RABUG2BA	L01081 Oryctolagus
28	166.2	23.2	1753	88	AF016310	AF016310 Macaca fa
29	166	23.2	1662	89	AF294901	AF294901 Macaca mu
30	166	23.2	1768	89	AF112113	AF112113 Macaca fa
31	166	23.2	2648	97	MFU91582	U91582 Macaca fasc
32	164.6	23.0	1088	95	RATUDPA01	M35202 Rat UDP glu
33	164.6	23.0	1961	95	RATUDPGT	M13506 Rat liver U
34	164.4	23.0	1021	89	AF19876	AF19876 Homo sapi
35	164.4	23.0	139015	69	AC025488	AC025488 Homo sapi
36	164	22.9	1543	89	AF177273	AF177273 Homo sapi
37	164	22.9	1851	89	AF177272	AF177272 Homo sapi
38	158.4	22.2	2818	7	OCU72742	U72742 Oryctolagus
39	157.8	22.1	1858	94	MMUDPGT	X06358 Mouse mRNA
40	156.6	21.9	2799	97	HSUGT2BIO	X06359 H. sapiens U
41	156.6	21.9	138665	72	AC055794	AC055794 Homo sapi
42	156.6	21.9	212904	66	AC021146	AC021146 Homo sapi
43	154.2	21.6	1846	95	RNUDPTGR	X03478 Rat liver m
44	153.6	21.5	897	89	AF179877	AF179877 Homo sapi
45	153	21.4	1844	95	RNU06273	U06273 Rattus norv

## ALIGNMENTS

RESULT 1  
AC021146/c  
LOCUS  
DEFINITION Homo sapiens chromosome 4 clone RP11-468N14, WORKING DRAFT  
AC021146  
AC021146.4 GI:8568861  
SEQUENCE. 24 unordered pieces.  
AC021146  
AC021146.4 GI:8568861  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
human.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
On Jun 16, 2000 this sequence version replaced gi:7344259.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NH0468N14  
----- Summary Statistics -----

## FEATURES

source  
Sequencing vector: M13: 888  
Sequencing vector: plasmid: 128  
Chemistry: Dye-primer ET; 88% of reads  
Chemistry: Dye-terminator Big Dye; 12% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 199729 bases at least Q40  
Consensus quality: 203731 bases at least Q30  
Consensus quality: 206340 bases at least Q20  
Insert size: 213000; agarose-fp  
Insert size: 210604; sum-of-ctligs  
Quality coverage: 3.75 in Q20 bases; sum-of-ctligs  
Quality coverage: 3.83 in Q20 bases; sum-of-ctligs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 24 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1564: contig of 1564 bp in length  
1 1565 1664: gap of unknown length  
1 1665 3085: contig of 1421 bp in length  
1 3086 3185: gap of unknown length  
1 3186 5134: contig of 1949 bp in length  
1 5135 5234: gap of unknown length  
1 5235 8169: contig of 2935 bp in length  
1 8170 8269: gap of unknown length  
1 8270 10341: contig of 2072 bp in length  
1 10342 10441: gap of unknown length  
1 10442 13614: contig of 3173 bp in length  
1 13615 13714: gap of unknown length  
1 13715 17089: contig of 3375 bp in length  
1 17090 17189: gap of unknown length  
1 17190 20701: contig of 3512 bp in length  
1 20702 20801: gap of unknown length  
1 20802 25001: contig of 4200 bp in length  
1 25002 25101: gap of unknown length  
1 25102 29020: contig of 3919 bp in length  
1 29021 29120: gap of unknown length  
1 29121 33356: contig of 4236 bp in length  
1 33357 33456: gap of unknown length  
1 33457 36902: contig of 3446 bp in length  
1 36903 37003: gap of unknown length  
1 37004 42975: contig of 5973 bp in length  
1 42976 43075: gap of unknown length  
1 43076 48339: contig of 5264 bp in length  
1 48340 48439: gap of unknown length  
1 48440 56329: contig of 8090 bp in length  
1 56330 56629: gap of unknown length  
1 56630 65213: contig of 8584 bp in length  
1 65214 65313: gap of unknown length  
1 65314 74715: contig of 9402 bp in length  
1 74716 74815: gap of unknown length  
1 74816 88546: contig of 13731 bp in length  
1 88547 88646: gap of unknown length  
1 88647 103367: contig of 14721 bp in length  
1 103368 103467: gap of unknown length  
1 103468 118167: contig of 14700 bp in length  
1 118168 118267: gap of unknown length  
1 118268 132765: contig of 14498 bp in length  
1 132766 132865: gap of unknown length  
1 132866 154092: contig of 21227 bp in length  
1 154093 154192: gap of unknown length  
1 154193 182144: contig of 27952 bp in length  
1 182145 182244: gap of unknown length  
1 182245 212904: contig of 30660 bp in length.  
Location/Qualifiers  
1. 212904  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="4"









-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 53 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 1640: contig of 1640 bp in length
1641 1740: gap of 100 bp
1741 2943: contig of 1203 bp in length
2944 3043: gap of 100 bp
3044 4897: contig of 1854 bp in length
4898 4997: gap of 100 bp
4998 6114: contig of 1117 bp in length
6115 6214: gap of 100 bp
6215 7665: contig of 1451 bp in length
7666 7765: gap of 100 bp
7766 9188: contig of 1423 bp in length
9189 9288: gap of 100 bp
9289 11371: contig of 2083 bp in length
11372 11471: gap of 100 bp
11472 13194: contig of 1723 bp in length
13195 13294: gap of 100 bp
13295 14324: contig of 1030 bp in length
14325 14424: gap of 100 bp
14425 15511: contig of 1087 bp in length
15512 15611: gap of 100 bp
15612 17362: contig of 1751 bp in length
17363 17462: gap of 100 bp
17463 19182: contig of 1720 bp in length
19183 19282: gap of 100 bp
19283 20717: contig of 1435 bp in length
20718 20817: gap of 100 bp
20818 22066: contig of 1249 bp in length
22067 22166: gap of 100 bp
22167 23790: contig of 1624 bp in length
23791 23890: gap of 100 bp
23891 25462: contig of 1572 bp in length
25463 25562: gap of 100 bp
25563 27562: contig of 2000 bp in length
27563 27662: gap of 100 bp
27663 29314: contig of 1652 bp in length
29315 29414: gap of 100 bp
29415 31317: contig of 1903 bp in length
31318 31417: gap of 100 bp
31418 33429: contig of 2012 bp in length
33430 33529: gap of 100 bp
33530 35133: contig of 1604 bp in length
35134 35233: gap of 100 bp
35234 37156: contig of 1923 bp in length
37157 37256: gap of 100 bp
37257 39661: contig of 2405 bp in length
39662 39761: gap of 100 bp
39762 41496: contig of 1735 bp in length
41497 41596: gap of 100 bp
41597 43309: contig of 1713 bp in length
43310 43409: gap of 100 bp
43410 45445: contig of 2036 bp in length
45446 45545: gap of 100 bp
45546 47118: contig of 1573 bp in length
47119 47218: gap of 100 bp
47219 49435: contig of 2217 bp in length
49436 49535: gap of 100 bp
49536 52546: contig of 3011 bp in length
52547 52646: gap of 100 bp
52647 53529: contig of 883 bp in length
53530 53629: gap of 100 bp
53630 55789: contig of 2160 bp in length
55790 55889: gap of 100 bp
55890 59029: contig of 3140 bp in length
59030 59129: gap of 100 bp

```

```

59130 61713: contig of 2584 bp in length
61714 61813: gap of 100 bp
61814 63696: contig of 1883 bp in length
63697 63796: gap of 100 bp
63797 65790: contig of 1994 bp in length
65791 65890: gap of 100 bp
65891 67845: contig of 1955 bp in length
67846 67945: gap of 100 bp
67946 69933: contig of 1988 bp in length
69934 70033: gap of 100 bp
70034 72807: contig of 2774 bp in length
72808 72907: gap of 100 bp
72908 75815: contig of 2908 bp in length
75816 75915: gap of 100 bp
75916 79940: contig of 4025 bp in length
79941 80040: gap of 100 bp
80041 83640: contig of 3600 bp in length
83641 83740: gap of 100 bp
83741 86921: contig of 3181 bp in length
86922 87021: gap of 100 bp
87022 91044: contig of 4023 bp in length
91045 91144: gap of 100 bp
91145 94259: contig of 3115 bp in length
94260 94359: gap of 100 bp
94360 96910: contig of 2551 bp in length
96911 97010: gap of 100 bp
97011 100209: contig of 3199 bp in length
100210 100309: gap of 100 bp
100310 105003: contig of 4694 bp in length
105004 105103: gap of 100 bp
105104 110031: contig of 4928 bp in length
110032 110131: gap of 100 bp
110132 115352: contig of 5221 bp in length
115353 115452: gap of 100 bp
115453 120961: contig of 5509 bp in length
120962 121061: gap of 100 bp
121062 126155: contig of 5094 bp in length
126156 126255: gap of 100 bp
126256 130511: contig of 4256 bp in length
130512 130611: gap of 100 bp
130612 138232: contig of 7621 bp in length.
Location/Qualifiers
1. .138232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-3C24"
/clone_11b="RPC1-11 Human Male BAC"
1. .1640
/note="assembly-fragment"
1741. 2943
misc_feature
/note="assembly-fragment"
3044. 4897
misc_feature
/note="assembly-fragment"
4998. 6114
misc_feature
/note="assembly-fragment"
6215. 7665
misc_feature
/note="assembly-fragment"
7766. 9188
misc_feature
/note="assembly-fragment"
9289. 11371
misc_feature
/note="assembly-fragment"
11472. 13194
misc_feature
/note="assembly-fragment"
13295. 14324
misc_feature
/note="assembly-fragment"
14425. 15511
misc_feature
/note="assembly-fragment"
15612. 17362
misc_feature
/note="assembly-fragment"
17463. 19182
misc_feature
/note="assembly-fragment"
19283. 20717
misc_feature
/note="assembly-fragment"

```

```

misc_feature      20818..22066
                  /note="assembly_fragment"
misc_feature      22167..23790
                  /note="assembly_fragment"
misc_feature      23891..25462
                  /note="assembly_fragment"
misc_feature      25563..27562
                  /note="assembly_fragment"
misc_feature      27663..29314
                  /note="assembly_fragment"

Query Match      56.8%; Score 406; DB 63; Length 138232;
Best Local Similarity 76.3%; Pred. No. 9,6e-93;
Matches 542; Conservative 0; Mismatches 155; Indels 13; Gaps 3;

QY      1 atgagatctgaacagtcagcttctgattctctctccagcagctctctgcttgcgt 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101021 AAGGTGCTGAGAGAGGGGTTTTCACAAATTCCTACCTCAGCTCTGCTGCTGCTGT 100962

QY      61 ggattctggggaagagctctgtgtgtgcccctgtgacatgagccattgcttaatgcaag 120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100961 GGAATCTTGAGAGAGGTAAGTGGGAGAGCACTGTAACATCAGCCATTGGCTTAATGTCAC 100902

QY      121 gtcattctgaagagctcatgtaagagagccatgagtaacagattgactcactca-aa 179
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100901 AATATTTTGGAGAACTCATAGAGAGGGGTCAACAGTAACACTGTTGACTCACACAC 100842

QY      180 gccctctgaattgactacagagaagccctctcattgaaattgaggtgtccatagcc 239
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100841 AGCGTTTCACTGACTAGAGGAGAGCCCTCTGCATTCGAACTTGAGCTGCTCCAGTGC 100782

QY      240 acagagcaga-----acagagaagaataatattgttgacctagctctgagat 290
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100781 ACATGACAAAGTGAATTTTGACATATGATTAATGAAATTTCTGACCTTACTGTAATGT 100722

QY      291 ctgtcagagcttcaacactgagcaatcagcttataaataaagattttgttgaat 350
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100721 CTGTGCAAGCTTGACCTCCCTGGCAGTCAAGAAAGAACTGAATATTTT---CTGATAC 100665

QY      351 aagaggaactttaaaatgatgtgtgagagcttattctacacagacactatgaaga 410
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100664 TAGAGAAATTTTAAACCTCTGTGTAGAGTTTGGCTACAAATCAGACACTCTTTGAGAT 100605

QY      411 gctacaggaacacactagatgataagccttataagccctgtagtcccgtagagacct 470
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100604 GCAAGAGGAACCCCACTAAACTGTAGCTTATAGACTGCTGCCATTCCTGTGGGAGTT 100545

QY      471 gatgctgagctgtctcagctcccttctgtctcacttagaactctcctaagagc 530
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100544 TATTGCTGAGTCTCTGCAATTCCTTTTGTGTCACAACTTAGAGCTTCTGTAGCTGGCAA 100485

QY      531 tatgagcgaagctgtgtggaacttccagctcacttctctatgtaacgtgtgctatgac 590
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100484 TATTGAGTGAAGCTCTGGGAAACTCTCACCTCCACTTCTATGTATCTGTGCTTAAGAC 100425

QY      591 aggaactaacagagaatgaccttctgaaagagtaaaaaatcaatgcttccagttt 650
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100424 AGGACATACGGAATTAAGACTTTCTTGGAAGAAAGTAAATTAACAATGCTGTGCATCTT 100365

QY      651 gtccacttctgattcaagattacgactatcattttgggaagagttt 700
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100364 TTTTAACCTTGCTGCCTCCAGATTTTAACTTTTGAATTTTGGACATTTT 100315

RESULT      6
LOCUS      AC013296 226077 bp DNA HTG 13-JUL-2000
DEFINITION Homo sapiens clone Rpl1-3M18, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC013296
VERSION AC013296.5 GI:9121251
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 226077)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone Rpl1-3M18
Unpublished
2 (bases 1 to 226077)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckery,L., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., DeRellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Headford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kahn,L., Karatas,A., Klein,J.,
Lewoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McMurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollard,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6514003.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L2775
Center clone name: 3_M_18

* NOTE: This record contains 246 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1
* 815 914: gap of 100 bp
* 915 1704: contig of 790 bp in length
* 1705 1804: gap of 100 bp
* 1805 2602: contig of 798 bp in length
* 2603 2702: gap of 100 bp
* 2703 3484: contig of 782 bp in length
* 3485 3584: gap of 100 bp
* 3585 4378: contig of 794 bp in length
* 4379 4478: gap of 100 bp
* 4479 5257: contig of 779 bp in length
* 5258 5357: gap of 100 bp
* 5358 6150: contig of 793 bp in length
* 6151 6250: gap of 100 bp
* 6251 7038: contig of 788 bp in length
* 7039 7138: gap of 100 bp
* 7139 7940: contig of 802 bp in length
* 7941 8040: gap of 100 bp
* 8041 8869: contig of 829 bp in length
* 8870 8969: gap of 100 bp
* 8970 9790: contig of 821 bp in length
* 9791 9890: gap of 100 bp
* 9891 10695: contig of 805 bp in length
* 10696 10795: gap of 100 bp
* 10796 11592: contig of 797 bp in length

```





[illegible]



NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Caranese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://Bacpac.med.buffalo.edu>)

#### VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is CTD-2005D20. Actual start of this clone is at base position 1 of RP11-185H6; actual end is at base position 181117 of RP11-185H6.

#### FEATURES

##### Source

Location/Qualifiers

1. 181117

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-185H6"

/clone\_lib="RPCI-11"

1. 71

/rpt\_family="ERV1"

271. 444

/rpt\_family="ERV1"

472. 592

/rpt\_family="ERV1"

593. 938

/rpt\_family="MaLR"

939. 1217

/rpt\_family="ERV1"

1617. 3178

/rpt\_family="L1"

3189. 5532

/rpt\_family="L1"

5533. 5818

/rpt\_family="Alu"

5819. 6208

/rpt\_family="L1"

6219. 6642

/rpt\_family="ERV1"

6660. 7156

/rpt\_family="ERV1"

9336. 10141

/rpt\_family="ERV1"

10149. 10219

/rpt\_family="L1"

10242. 10508

/rpt\_family="ERV1"

10509. 10978

/rpt\_family="L1"

10979. 11006

/rpt\_family="ERV1"

repeat\_region 10988. 12442  
/rpt\_family="ERV1"  
repeat\_region 12446. 12879  
/rpt\_family="ERV1"  
repeat\_region 12921. 13150  
/rpt\_family="L1"  
repeat\_region 13187. 13499  
/rpt\_family="L1"  
repeat\_region 13549. 13976  
/rpt\_family="L1"  
repeat\_region 14032. 14404  
/rpt\_family="L1"  
repeat\_region 14499. 14598  
/rpt\_family="ERV1"  
repeat\_region 14628. 16122  
/rpt\_family="L1"  
repeat\_region 16135. 16537  
/rpt\_family="ERV1"  
repeat\_region 16615. 16996  
/rpt\_family="L1"  
repeat\_region 16998. 17293  
/rpt\_family="Alu"  
repeat\_region 17710. 17766  
/rpt\_family="Alu"  
repeat\_region 17801. 18334  
/rpt\_family="L1"  
repeat\_region 18343. 18622  
/rpt\_family="L1"  
repeat\_region 18623. 19027  
/rpt\_family="MaLR"  
repeat\_region 19028. 20370  
/rpt\_family="L1"  
repeat\_region 20386. 20698  
/rpt\_family="ERV1"  
repeat\_region 20699. 20938  
/rpt\_family="ERV1"  
repeat\_region 20977. 21022  
/rpt\_family="ERV1"  
repeat\_region 21043. 22020  
/rpt\_family="ERV1"  
repeat\_region 22022. 22314  
/rpt\_family="ERV1"  
repeat\_region 22321. 22476  
/rpt\_family="L1"  
repeat\_region 22500. 23925  
/rpt\_family="L1"  
repeat\_region 24128. 24222  
/rpt\_family="ERV1"  
repeat\_region 24452. 24625  
/rpt\_family="ERV1"  
repeat\_region 24645. 24899  
/rpt\_family="L1"  
repeat\_region 24916. 25128  
/rpt\_family="MaLR"  
repeat\_region 25129. 25436  
/rpt\_family="Alu"  
repeat\_region 25437. 25499  
/rpt\_family="MaLR"  
repeat\_region 27081. 27433  
/rpt\_family="MaLR"  
repeat\_region 27437. 28027  
/rpt\_family="L1"  
repeat\_region 28021. 28967  
/rpt\_family="L1"  
repeat\_region 28954. 29769  
/rpt\_family="L1"  
repeat\_region 29770. 30074  
/rpt\_family="Alu"  
repeat\_region 30075. 30629  
/rpt\_family="L1"  
repeat\_region 30631. 31554  
/rpt\_family="L1"  
repeat\_region 31675. 32162





```

RESULT 12
AF112112 2108 bp mRNA PRI 11-APR-1999
LOCUS Macaca fascicularis UDP-glucuronosyltransferase 2B19 precursor,
DEFINITION mRNA, complete cds.
ACCESSION AF112112 GI:4580601
VERSION AF112112.1
KEYWORDS crab-eating macaque.
SOURCE Macaca fascicularis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 2108)
AUTHORS Belanger, G., Barbier, O., Hum, D. W. and Belanger, A.
TITLE Molecular cloning, expression and characterization of a monkey
steroid UDP-glucuronosyltransferase, UGT2B19, that conjugates
testosterone
JOURNAL Eur. J. Biochem. 260 (3), 701-708 (1999)
MEDLINE 99203465
REFERENCE 2 (bases 1 to 2108)
AUTHORS Belanger, G., Barbier, O., Hum, D. W. and Belanger, A.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-1998) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Sainte Foy, Quebec G1V 4G2, Canada
FEATURES
Source Location/Qualifiers
1..2108
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
26..1612
/EC_number="2.4.1.17"
/note="UGT2B19"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B19 precursor"
protein_id="AAD24435.1"
/db_xref="GI:4580602"
/translation="MSMKWTSAALLLIQLSCYLSFGSGKVLVMPFEFSHMNIKRTILD
ELVORGHEVTLAVSTSLPDPNPSPLKFEICPSLTPEFODSVTOLVKRMSDRK
DPEHPFLHVOEMMTYGDIMIKKCKDYVSNKKMKLOESRFDVLAIAISPGEGL
AELKIPVYSLRSPFGTALKEHGGFELPSPVPTVSLKLDQMTEFMRVONMTYV
YDFEWQVVDVKNWDFQYKVLGRPTLLEIMAKAIVLIRNWFQFHPILPNVEF
VGLRCKPAKPLPKEMEERFVOSGDNVGVFSLGSVSNMSEERANVIAIALKIPQK
VLMRDGNKPDITGLNTOLYKWLPMNDLGHKPTRAFTHGGANGYEALVIGIPWG
VLPADDPNINAHMKAKGAVALDPTMSTPLNALKIVINDPIKEMAMLSITH
DQPVKPLDRAVWIEFVNRHKAHKAHLRAVAHDLTPQYHSLDYIGFLACVATVIFII
TCLFCVMKFEVTRKKGRKD"
BASE COUNT 649 a 393 c 438 g 628 t
ORIGIN
Query Match 26.3%; Score 188.4; DB 89; Length 2108;
Best Local Similarity 55.8%; Pred. No. 1.4e-37;
Matches 404; Conservative 0; Mismatches 311; Indels 9; Gaps 2;
QY 1 atgaagctcgacaagtcagcctgtgtatcttcgtctcgtcgaagctcttcgtt--gttggc 57
DB 23 AGGATGTCTATGAAATGAGACTTCAGCTCTTCTCTCATACAGCTGAGCTGTACTTGAGC 82
QY 58 tgtgaattctgtggaaagctcgtgtgtgcccgtgtgacatgacatgacatgcttaatgtc 117
DB 83 TTTGGAGATTGTGAAAGGCTGTGCTGTGGCCACAGAAATTCAGCATTCGATGAATATA 142
QY 118 aaggtatcttagaagaagctcatagtagagagagcagtagaagtagaagtagaagtaga 177
DB 143 AAGACAAATCTCTGATGATGCTTGTCCAGAGAGTCATGAGGTGACTGACTGACATATCA 202
QY 178 aagccttcgttaattgactacagaagccttcgacttgaattgaattgagtggtccatag 237
DB 203 ACTTCATTTCTTCTCTGATCCCAACACCCATCTCTTAATTTGAAATTTGTCTCTACA 262
QY 238 ccacagagacagacagaagaataatgtgtgacctagctcga-----atgtc 291

```

```

DB 263 TCCTTAACCGAAGCTGATTTCAGATAGCGTCACGCCAACCTGGTTAAGAGATGTCAGAC 322
QY 292 ttgcagagctatcaaccctggcaatcagttataaataatgattttttgttgaata 351
DB 323 ATTGGAAGACACATTTTGGCCACATTTTTCACAGCTACAGAAATGATGTGCACATAT 382
QY 352 agagaactttaaaatgtagtgtgagagcttatactacatcagacactatgaagaag 411
DB 383 GGTGACATGATTTAGAAAGTCTCTGAAGATGCTGTTCTATTAAGAACTTATGAAGAA 442
QY 412 ctacaggaacacacactagatgtgaacgcttatagaccctgttgatcccgtagagac 471
DB 443 CTACAGAGACCAAGATTTGATGTTGTTGCTGCGAGATGCTATTCGCCCTGTGAGCTA 502
QY 472 atgctgagttgttccagtcagccttcttgcttaacacttgaactctcctcaaggcaat 531
DB 503 CTGCTGAGCTACTTAAATAACCTTGTGTATACAGCTCCATTTCTCTCTGCTGACGA 562
QY 532 atgagcgaagctgtggaacactccagctccacttccctgtactcgtgacctagaca 591
DB 563 CTGGAAGACATGATGTGAGAGGTTTCTGTCCTCCCTTCCATGACTGTACTATGTCA 622
QY 592 ggaactacagacagaatgaccttcttgaaagaagtaaaaatcaatgcttcaagtttg 651
DB 623 GAATTAAGAGATCAATATGACTTTTCATGAGAGAGTACAAATATGATCTATATGTTAT 682
QY 652 ttccacttctggaattcaagattacgatacatcttcttggaagaagttatagtagca 711
DB 683 TTTGACTTTTGTTCAGATGATGATGTAAGAACTGGATTCAGTTTACAGTAAAGTT 742
QY 712 ttag 715
DB 743 CTAG 746
RESULT 13
HSAJ5162
LOCUS Homo sapiens mRNA for UDP-glucuronosyltransferase. 24-JUN-1998
DEFINITION AJ005162.1 GI:3135024
ACCESSION AJ005162.1
VERSION AJ005162.1
KEYWORDS UDP-glucuronosyltransferase; UGT2B4 gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2079)
AUTHORS Rittler, J.K.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-1998) Rittler, J.K., Pharmacology and Toxicology,
of Virginia, Commonwealth University, P.O. Box 980613, Medical College
of Virginia, Richmond, Virginia, 23113-0613, USA
2 (bases 1 to 2079)
AUTHORS Rittler, J.K., Chen, F., Sheen, Y.Y., Lubets, R.A. and Owens, I.S.
TITLE Two human liver cDNAs encode UDP-glucuronosyltransferases with 2
log differences in activity toward parallel substrates including
hydroxycholeic acid and certain estrogen derivatives
JOURNAL Biochemistry 31 (13), 3409-3414 (1992)
MEDLINE 92207964
REFERENCE 3 (bases 1 to 2079)
AUTHORS Jackson, M.R., McCarthy, L.R., Harding, B., Wilson, S., Coughtrie, M.W.
and Burchell, B.
TITLE Cloning of a human liver microsomal UDP-glucuronosyltransferase
JOURNAL CDNA Biochem. J. 242 (2), 581-588 (1987)
MEDLINE 87241362
FEATURES
Source Location/Qualifiers
1..2079
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UDP2B4-1"
/clone_11b="Lambda bacteriophage gt11"
/issue_11b="Liver"

```

```

gene      34. .1620
          /gene="UGT2B4"
CDS       34. .1620
          /gene="UGT2B4"
          /EC_number="2.4.1.17"
          /function="transfers glucuronic acid to hydroxycholeic
          acid and estrogen catechols"
          /codon_start=1
          /evidence=experimental
          /product="UDP-glucuronosyltransferase"
          /protein_id="CAA06396.1"
          /db_xref="GI:3135025"
          /db_xref="SPTREMBL:O60867"
          /translacion="MSMKWTSAALLIOLSCYFSSGCKVLMWPTFESHMNIKTILD
          ELVORGEVTVLASSASISFDPNSPTLKFEYVPSVLTKEFEDIIKOLVMAELPK
          DFNMSFGEVTVLASSASISFDPNSPTLKFEYVPSVLTKEFEDIIKOLVMAELPK
          AELKIPFYYS.LRPSGVA.IEKHSGGIL.FPSPVYVMSLSDMTFERVNMNYVL
          YEFWFOIPDMKRMDOFSEVLRPTTSETYAKADI.IRMVMPORPHILPVEF
          VGLHCKPAKPLPKEMEFEVSSGSENGVVFSLGSVNSTSERANVITASALAKIPK
          VLMRDGNKPDILGLNTRLYKWI.PONDLLGHPKTRAFITTHGANGIYEAIYHGTIPWG
          VPLEADOPNIAHMKKGAAYSLDEFHTMSDILNALKTVINDPYLEKEMAKLSRHH
          DDPVRLDRAVMEFEVMSKGAHLRVAHADLTFWQYHSLDTGTFLLACVAVTIFII
          TKLCFVWKFVRKGRKRD"
          346
conflict  /gene="UGT2B4"
          /citalion="[3]
          /replace="g"
          548. .549
conflict  /gene="UGT2B4"
          /citalion="[3]
          /replace="ct"
          1187. .1188
conflict  /gene="UGT2B4"
          /citalion="[3]
          /replace="ct"
          1190
conflict  /gene="UGT2B4"
          /citalion="[3]
          /replace="c"
          1192
conflict  /gene="UGT2B4"
          /citalion="[3]
          /replace="a"
          1977. .1978
conflict  /citalion="[3]
          /replace="cga"
          2068
          /citalion="[3]
          /replace="g"
BASE COUNT 634 a 396 c 436 g 613 t
ORIGIN
Query Match 26.1%; Score 186.8; DB 93; Length 2079;
Best Local Similarity 55.7%; Pred. No. 3.6e-37;
Matches 403; Conservative 0; Mismatches 312; Indels 9; Gaps 2;
QY 1 atgaggtctgacagtcagcttgattcttcgtcccgagctctctctgt---gttgc 57
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 31 AGGATGCTCTATGAAATGACCTTACCTCTGCTGATACACCTGAGCTGTTACTTTAC 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 58 tctgagttctgtggaagctccctgtgtgcccctgtgacatgagcatgtgcttaatg 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 91 TCTGGAGTTGTGGAAAGTGTGCTGTGTGGCCACAGATTCAGCCACTGATGATATA 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 aaggtcatctagaagagctcatagtgagagagccatgagtaacagatgactcaacta 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 151 AAGACAATCTGATGATGATGCTGCTGAGAGAGGTCATGAGTGTATTTGGCATCTTCA 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 178 aagccttgatgaatgactacaggaagcctctgcatgaattgaggtgtgcatatg 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 211 GCTTCATTTCTTTGATGCCACAGCCCATCTACTTAAATTGAATTATCTCTGA 270

```

```

QY 238 ccacagacagacagagaagaataatgttctgacacagctctga-----atgic 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 TCTTAACTAAACTGAGTTTGAGCATATTTATCAAGCAGCTGCTTAGAGTGGCAGAA 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 ttgcaggtctatcaacctggcaatcagttataaataatgatttttctgtgaata 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 CTTCAGAAAGACATTTTGTGTCATATTTTTCACACAGTACAGAAATCATGTGACATTT 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 agaggaacttaaaaatgagtgtgagagctcttatctacatctgacacattatgaag 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 391 AATGACATACCTTAGAAAGTCTCTGTAAGCATATAGTTTCAATTAAGAAACTATAGAAA 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 ctacaggaacacacactacagatgtaacgttataagaccctgtgaltcccggtgagact 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 451 CTACAGGACTCAAGTTTGAATGTTGTTCTTGCAGATGCTCTTTCCCTTTGGTGAGCTG 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 472 atgctgagttgtctcagctccctcttctgtctcaacttgaactctctaagagagca 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 511 CTGGCGGAGTTACTTAAATACCTTTGTCATACACCTGCTGCTCTGCTGACGCA 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 532 atgagagcgaagctgtggaagaaacttcagctccacttccatgtacctgtgctatgaca 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 571 ATTTGAAAGACATAGTGAGAGACTTCTGTCCTCTCCATATGCTCTGTATATGCA 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 592 ggactaacagacagaaatgaccttctctggaagaagaanaatcaatgcttcaagtttg 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 631 GAACATGAGTACCAATGACTTTTATAGAGAGTAAAGAAATATGATGTATGCTTAT 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 ttccactctgattcagatgaatgaactatcattttgttgaagtttatagtaagca 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 691 TTTGAATTTTGTGTTCCAAATATTTTGACATGAAGAACTGGATCACTTACAGTAACT 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 712 tttag 715
    ||| |||
DB 751 CTAG 754

```

RESULT 14

HSUBPGT 2093 bp mRNA PRI 23-MAR-1995

LOCUS Human mRNA for liver microsomal UDP-glucuronosyltransferase

DEFINITION (UDPGT).

ACCESSION Y00317

VERSION Y00317.1 GI:37588

KEYWORDS transferase; UDP-glucuronosyltransferase.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2093)

AUTHORS Burchell, B.

TITLE Direct Submission

REFERENCE Submitted (26-JAN-1987) Burchell B., Dept. Biochemistry, Medical Sciences Institute, University Dundee, Scotland

2 (bases 1 to 2093)

JACKSON, M.R., MCCARTHY, L.R., HARDING, J., WILSON, S., COUGHLIN, M.W. and BURCHELL, B.

Cloning of a human liver microsomal UDP-glucuronosyltransferase cDNA

Biochem. J. 242 (2), 581-588 (1987)

FEATURES

source

1. .2093

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/tissue\_type="liver"

/clone\_id="lambda gt11 liver cDNA"

/clone="Hug25"

38. .106

/product="put. signal peptide"

38. .1624

/note="precursor"

/codon\_start=1



mat_peptide	107. .1621	/product="nature UDPGT (AA 1-505) (EC 2.4.1.17)"
misc_feature	979. .988	/note="pot. glycosylation site"
misc_feature	2076. .2081	/note="pot. polyA signal"
BASE COUNT	638 a 397 c 439 g 619 t	
ORIGIN	107. .1621	
	TKCLCFWKAFVPTGKKKKRD"	
	DQPKPLDRAVFTIEFYMRHKGAKHLVAADLTWFOYHSLDTGFLACVATYIFIT	
	ELVARGHEVTVIASSASISFDPNPSPTLKFVYVSLTKTFEEDI IKOLVYRAEELPK	
	DTFWSFSVOEIMWMTFNDLIRKCKDIVSKIPVVKMLQESLSDQMTIERKNNITYV	
	AELKLPFYVTRPRSPGATIEKHSGGLTPEPSYIPVVMSELSRDPDITIERKNNITYV	
	YEEFWFIIDPMKKWDOFYSEVLGRPTLLSTFMAADITWLIRNYWDFQPHLLPNVEE	
	VGGJLHCKPAKPLPKEMEEFVQSSGCEVNVVSLCSMYSNTESEERANYIASLALIPOR	
	VLAFEDCNKPDPTGLTLRLYKWIPONDLIGPKPRATFTHGAGANGYIKASLPRIPMVG	
	VPLPADPDINAHMKAKAGCAVSLIDFHMSSDILNATKATYINDPLYENEMAKSLRIHH	
	/protein_id="CA68415.1"	
	/db_xref="GI:37589"	
	/db_xref="SWISS-PROT:P06133"	
	/translation="MSKMTKTSALLLIQLSCYFSSGCKVLVWPTFESHMNIKITILD	
	Query Match 26.1%; Score 186.8; DB 97; Length 2093;	
	Best Local Similarity 75.7%; Pred. No. 3.6e-37;	
	Matches 403; Conservative 0; Mismatches 312; Indels 9; Gaps 2;	

Db	695	TTTGAATTTTGGTTCCCAATATTTCAGTCATGAGGAGAGTGGTTCACGTTCTACAGTGAAGTT	754
Qy	712	ttag 715 	
Db	755	CTAG 758	
RESULT	15		
AF064200			
LOCUS			
DEFINITION	AF064200	2097 bp mRNA	PRI 11-DEC-1998
ACCESSION		Homo sapiens UDP-glucuronosyltransferase 2B4 precursor (UGT2B4)	
VERSION		NRNA, UGT2B4*E458 allele, complete cds.	
KEYWORDS		AF064200.1 GI:3153831	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
TITLE		1 (bases 1 to 2097)	
JOURNAL		Levesque, E., Beaulieu, M., Hum, M. and Belanger, A.	
REFERENCE		Characterization and substrate specificity of UGT2B4(E458) : A	
AUTHORS		UDP-glucuronosyltransferase encoded by a polymorphic gene	
		Pharmacogenetics (1999) In press	
		2 (bases 1 to 2097)	
		Levesque, E., Beaulieu, M. and Belanger, A.	

OY	1	atgagctgcagaagtcacgcttctgattcttcgtccctgcagcctctcttcttgcg	57
Db	35	AGGATGTCATATGAANTGAGACTTCAGCTCTTCTGCTGATACAGCTGAGACTGTACTTACG	94
OY	58	tctgattctctgtggaaagtcctcgtgctgtgagccctgtacatlgagccatctgcttaatgctc	117
Db	95	TCTGGAGATTGTGGAAAGTGCTCTGTGTGGCCCAACAATTCACCACTGATGATCAATA	154
OY	118	aagtcattctcagaagagcgtcatatgtagagggccatgaggttaaaagcatttgactcaetca	177
Db	155	AAGACATCCCGATGTAACCTTGTCACAGAGGTCATAGGAGTATGATGGCATCTTCA	214
OY	178	aagccttcgttaattgcatacagaagcctctcgtcatlgaaatttggagtggtcccatatg	237
Db	215	GCTTCCATTTCTTTCGATCCCAACAGCCCATCTACTTTAAATTTGAAAGTTTATCCTGTA	274
OY	238	ccacaagacagacagagaanaatlgaaatalctgttgacctagctctga-----atgctc	291
Db	275	TCTTTAACTAAATCTGATGATTGAGGATATTATCAAGCAGCTGGTTAGAGATGGCGAGA	334
OY	292	ttggcagcgttatcaacctggtgcaatcagttataaataatgatgttttttgttgaata	351
Db	335	CTTCCAAAAGACACATTTTGGTCATATTTTTCACAACTACAAAGAAATCAATGTGACATTTT	394
OY	352	agagaacctttaaanaatgatgtgtgtgagagcttatctacaatcagacacataigaagaag	411
Db	395	AATGCACTACTTACAAAGTTCTGTAAAGATATATGCTTTCAAATAAGAAACTTATACAGAAA	454
OY	412	ctacaagaaacaaactacgaatgtaacgcttatagaacctgtgattcccgctlgagagacctg	471
Db	455	CTACAGGAGTCACAGATTGTGATGTTGTTCTTTCACAGATGCTGTTTTCCCTTTGGTGACTG	514
OY	472	atgagctgagttgcttccaagtcctcttctgtgctcaacttgaacttctctaagaagcgaat	531
Db	515	CTGGCCGAGTACTTAAATACCTTTGTCTACAGGCTCGCTTCCTCTCCGTGGCTACGCA	574
OY	532	atgagagcgaagcctgttggaanaactcccaagtcctccacttccatagcactgtctgcatacaga	591
Db	575	ATTGAAAAGCATATGTGAGGCACTTCTGTCCCTCTTCTCTATGTGCCGTGTGTTATGTCA	634
OY	592	ggaactaacaagacgaatgaccttctctggaagaagtaaaaaatloaaatgcgtcttcagtttg	651
Db	635	GAACTAAAGTACCAATGACTTTTCATAGAGAGGGGTAAAAAATATGATCATATGTGCTTTAT	694
OY	652	ttccaactcttgattcagagattacgactatacatlttttggaagaagtttatagtaagca	711

[illegible]

OY	178	aagcttcgtaaatgtgctaaggaagccctgcgcatgaatcttaagtgctgcaatg	237
Db	203	GCCTTCATTTCTTTGCGATCCCAACGCCCATCTACTCTTAAATTGAAGTTTATCCTGTGA	262
OY	238	ccacaggacacagacaagaanaatgaaatacttgctgaccgaactctga----	atgic 291
Db	263	TCTTTAACTAAACTGAGTTTGAGAGATATTATCAAGACGCTGTTAAGAGATGGCAGAA	322
OY	292	ttgcacggctatcaacctggcaatcagttataaataatgaatctttttgttgtaata	351
Db	323	CTTCCAAAAGACACATTTTGGTCATATTTTTCACAACTACAAAGAAATCAATGACATT	382
OY	352	agaggaactttaanaatgagtggtgagagctttatctacatgaagcaactatgaaag	411
Db	383	AATGACATATCTTAAAGATGTTGTGAAGGATATAGCTTTCAATTAAGAACTTATGAAGAAA	442
OY	412	ctacaggaacacaactcagatgtaagcctatagaacctgtgactcccgctgagaacctg	471
Db	443	CTACAGGAGTCAAGATTTTGATGTTGTTCTTGCAAGATCGTGTTCCTTCCCTTGGTGAGCTG	502
OY	472	atgagctgagttgcttccagtcagtcctttgtgctcaacttgaacttctcctaagagcaat	531
Db	503	CTGGCGGAGTTACTTAAATATCCCTTTGTACAGCCCTCGCTTCCTCCGTGCTACGCA	562
OY	532	atgagcggaagctcttgggaaaactcccaagccaccactttccatgactgtgctcatgaca	591
Db	563	ATTGAAAAGCATATGTGGAGGAGCTCTGTGTCCTCCTTCATATGCGCTGTGTTATGTCA	622
OY	592	ggactacaagacagaatgaccttctctgaaaagtaaaaaatccaatgcttccagtttg	651
Db	623	GAACTATAGGACCAATGACTTTCATAGAGAGGGCTAAAAAAATATGATCTATGCTGTAT	682
OY	652	ttcaccactctgattccagatataagacatacatllttgggaagagtttataagaca	711
Db	683	TTTGAATTTTGGTTCCAATATTTTGACATGGAAGATGGGATCGTCTACAGTGAAGTT	742
OY	712	ttag 715	
Db	743	CTAG 746	

Search completed: August 27, 2001, 18:10:18  
Job time: 9691 sec